

## SEQUENCE LISTING

### (1) GENERAL INFORMATION:

#### (i) APPLICANT:

5 (A) NAME: Arch Development Corporation  
(B) STREET: 1101 East 58th Street  
(C) CITY: Chicago  
(D) STATE: Illinois  
(E) COUNTRY: USA  
10 (F) POSTAL CODE (ZIP): 60637

(A) NAME: Michael S. Parmacek  
(B) STREET: 1225 E. 56th Street  
(C) CITY: Chicago  
15 (D) STATE: IL  
(E) COUNTRY: USA  
(F) POSTAL CODE (ZIP): 60637

(A) NAME: Julian Solway  
(B) STREET: 746 Grove Street  
(C) CITY: Glencoe  
(D) STATE: IL  
(E) COUNTRY: USA  
20 (F) POSTAL CODE (ZIP): 60022

25 (ii) TITLE OF INVENTION: PROMOTER FOR SMOOTH MUSCLE CELL EXPRESSION

(iii) NUMBER OF SEQUENCES: 51

#### (iv) COMPUTER READABLE FORM:

30 (A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

#### (v) CURRENT APPLICATION DATA:

APPLICATION NUMBER: Unknown

#### (vi) PRIOR APPLICATION DATA:

40 (A) APPLICATION NUMBER: US 08/726,807  
(B) FILING DATE: 07-OCT-1996

### (2) INFORMATION FOR SEQ ID NO: 1:

#### (i) SEQUENCE CHARACTERISTICS:

45 (A) LENGTH: 1419 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GAATTCAGGA CGTAATCAGT GGCTGGAAAG CAAGAGCTCT AGAGGAGCTC CAGCTTATTA

	TGACCCTTCC	TTCAGATGCC	ACAAGGAGGT	GCTGGAGTTC	TATGCACCAA	TAGCTTAAAC	120
5	CAGCCAGGCT	GGCTGTAGTG	GATTGAGCGT	CTGAGGCTGC	ACCTCTCTGG	CCTGCAGCCA	180
	GTTCTGGGTG	AGACTGACCC	TGCTTGAGGG	TTCTCTCCTT	CCCTCTCTCT	ACTCCTTTCT	240
	CCCTCTCCCT	CTCCCTCTCT	CTGTTTCCTG	AGGTTTCCAG	GATTGGGGAT	GGGACTCAGA	300
10	GACACCACTA	AAGCCTTACC	TTTTAAGAAG	TTGCATTGAG	TGAGTGTGTG	AGACATAGCA	360
	CAGATAGGGG	CAGAGGAGAG	CTGGTTCTGT	CTCCACTGTG	TTTGGTCTTG	GGTACTGAAC	420
	TCAGACCATC	AGGTGTGATA	GCAGTTGTCT	TTAACCCTAA	CCCTGAGCCT	GTCTCACCTG	480
15	TCCCTTCCCA	AGACCACTGA	AGCTAGGTGC	AAGATAAGTG	GGGACCCTTT	CTGAGGTGGT	540
	AGGATCTTTC	ACGATAAGGA	CTATTTTGAA	GGGAGGGAGG	GTGACACTGT	CCTAGTCCTC	600
20	TTACCCTAGT	GTCTCCAGCC	TTGCCAGGCC	TTAAACATCC	GCCCATTTGC	ACCGCTCTAG	660
	AAGGGGCCAG	GGTTGACTTG	CTGCTAAACA	AGGCACTCCC	TAGAGAAGCA	CCCGCTAGAA	720
	GCATACCATA	CCTGTGGGCA	GGATGACCCA	TGTTCTGCCA	CGCACTTGGT	AGCCTTGGAA	780
25	AGGCCACTTT	GAACCTCAAT	TTTCTCAACT	GTTAAATGGG	GTGGTAACTG	CTATCTCATA	840
	ATAAAGGGGA	ACGTGAAAGG	AAGGCGTTTG	CATAGTGCCT	GGTTGTGCAG	CCAGGCTGCA	900
30	GTCAAGACTA	GTTCCACCA	ACTCGATTTT	AAAGCCTTGC	AAGAAGGTGG	CTTGTTTGTC	960
	CCTTGCAGGT	TCCTTTGTCT	GGCCAAACTC	TAGAATGCCT	CCCCCTTTCT	TTCTCATTGA	1020
	AGAGCAGACC	CAAGTCCGGG	TAACAAGGAA	GGGTTTCAGG	GTCCTGCCCA	TAAAAGGTTT	1080
35	TTCCCGGCCG	CCCTCAGCAC	CGCCCCGCC	CGACCCCCGC	AGCATCTCCA	AAGCATGCAG	1140
	AGAATGTCTC	CGGCTGCCCC	CGACAGACTG	CTCCAACCTG	GTGTCTTTCC	CCAAATATGG	1200
40	AGCCTGTGTG	GAGTGAGTGG	GGCGGCCCGG	GGTGGTGAGC	CAAGCAGACT	TCCATGGGCA	1260
	GGGAGGGGCG	CCAGCGGACG	GCAGAGGGGT	GACATCACTG	CCTAGGCGGC	CTTTAAACCC	1320
	CTCACCAGC	CGGCGCCCCA	GCCCGTCTGC	CCCAGCCCAG	ACACCGAAGC	TACTCTCCTT	1380
45	CCAGTCCACA	AACGACCAAG	CCTTGTAAGT	GCAAGTCAT			1419

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 991 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: CDS  
(B) LOCATION:38..218

(ix) FEATURE:

- (A) NAME/KEY: CDS  
(B) LOCATION:322..500

(ix) FEATURE:

- (A) NAME/KEY: CDS  
(B) LOCATION:866..967

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```
CTTTTCTCCA CACTCTATAC TTTAGCTCTG CCTCAAC ATG GCC AAC AAG GGT CCA      55
                                     Met Ala Asn Lys Gly Pro
                                     1           5

TCC TAC GGC ATG AGC CGA GAA GTG CAG TCC AAA ATT GAG AAG AAG TAT      103
Ser Tyr Gly Met Ser Arg Glu Val Gln Ser Lys Ile Glu Lys Lys Tyr
          10           15           20

GAC GAG GAG CTG GAG GAG CGA CTA GTG GAG TGG ATT GTA GTG CAG TGT      151
Asp Glu Glu Leu Glu Glu Arg Leu Val Glu Trp Ile Val Val Gln Cys
          25           30           35

GGC CCT GAT GTA GGC CGC CCA GAT CGT GGG CGC CTG GGC TTC CAG GTG      199
Gly Pro Asp Val Gly Arg Pro Asp Arg Gly Arg Leu Gly Phe Gln Val
          40           45           50

TGG CTG AAG AAT GGT GTG G TGAGTAACCC TTGCGAAGGG AATCTAGGGA      248
Trp Leu Lys Asn Gly Val
          55           60

TGTGTATGCC GCCCTACAAA CTGTGAGACA GACTCCCTGA GCTGAGTGTT CAGTTGTGTT      308

CTGTACCTGG CAG ATT CTG AGC AAA TTG GTG AAC AGC CTG TAT CCT GAG      357
          Ile Leu Ser Lys Leu Val Asn Ser Leu Tyr Pro Glu
          1           5           10

GGA TCG AAG CCA GTG AAG GTG CCT GAG AAC CCA CCC TCC ATG GTC TTT      405
Gly Ser Lys Pro Val Lys Val Pro Glu Asn Pro Pro Ser Met Val Phe
          15           20           25
```

35

(i) SEQUENCE CHARACTERISTICS:

- 40

- 45

50

Arg Leu Gly Phe Gln Val Trp Leu Lys Asn Gly Val  
50 55 60

(2) INFORMATION FOR SEQ ID NO: 4:

5 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 59 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

10 (ii) MOLECULE TYPE: protein  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Ile Leu Ser Lys Leu Val Asn Ser Leu Tyr Pro Glu Gly Ser Lys Pro  
1 5 10 15  
Val Lys Val Pro Glu Asn Pro Pro Ser Met Val Phe Lys Gln Met Glu  
20 25 30  
Gln Val Ala Gln Phe Leu Lys Ala Ala Glu Asp Tyr Gly Val Ile Lys  
20 35 40 45  
Thr Asp Met Phe Gln Thr Val Asp Leu Tyr Glu  
50 55

25 (2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 34 amino acids  
30 (B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

35 Lys Asp Met Ala Ala Val Gln Arg Thr Leu Met Ala Leu Gly Ser Leu  
1 5 10 15  
Ala Val Thr Lys Asn Asp Gly Asn Tyr Arg Gly Asp Pro Asn Trp Phe  
40 20 25 30  
Met Lys

45 (2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 575 base pairs  
50 (B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 28..169

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

```
ACTTACCCTG GTTCCTTTTC TTCTAGG AAA GCC CAG GAG CAT AAG AGG GAC      51
                      Lys Ala Gln Glu His Lys Arg Asp
                        1                      5

10 TTC ACA GAC AGC CAA CTG CAG GAG GGG AAG CAC GTC ATT GGC CTT CAA      99
   Phe Thr Asp Ser Gln Leu Gln Glu Gly Lys His Val Ile Gly Leu Gln
      10                15                20

15 ATG GGC AGC AAC AGA GGA GCC TCG CAG GCT GGC ATG ACA GGC TAT GGG      147
   Met Gly Ser Asn Arg Gly Ala Ser Gln Ala Gly Met Thr Gly Tyr Gly
      25                30                35                40

20 CGA CCC CGG CAG ATC ATC AGT T AGAAAGGGAA GGCCAGCCCT GAGCTGCAGC      199
   Arg Pro Arg Gln Ile Ile Ser
                        45

ATCCTGCTTA GCCTGCCTCA CAAATGCCTA TGTAGGTTCT TAGCCCTGAC AGCTCTGAGG      259

25 TGTCACTGGG CAAAGATGAC TGCACATGGG CAGCTCCAC CTATCCTTAG CCTCAGCCCA      319
   GCATCTTACC CCAGAGCCAC CACTGCCCTG GCCCCTGTTC CCAGCTGTAC CCCACCTCT      379

ACTGTTCTCTC TCATCCTGGA GTAAGCAGGG AGAAGTGGGC TGGGGTAGCT GGCTGTAGGC      439
30 CAGCCCACTG TCCTTGATAT CGAATGTCCT TTGAAGGAGA CCCAGCCCAG CCTCTACATC      499
   TTTTCCTGGA ATATGTTTTT GGGTTGAAAT TCAAAAAGGA AAAAAGAAAA ATATATAAAT      559

35 ATATATATAT ATATAC      575
```

(2) INFORMATION FOR SEQ ID NO: 7:

40 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 47 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

45 (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

```
Lys Ala Gln Glu His Lys Arg Asp Phe Thr Asp Ser Gln Leu Gln Glu
  1          5          10          15

50 Gly Lys His Val Ile Gly Leu Gln Met Gly Ser Asn Arg Gly Ala Ser
    20          25          30
```

Gln Ala Gly Met Thr Gly Tyr Gly Arg Pro Arg Gln Ile Ile Ser  
 35 40 45

5 (2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 1102 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ix) FEATURE:

- 15 (A) NAME/KEY: CDS  
 (B) LOCATION: 77..681

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

20 GCCCGTCTGC CCCAGCCCAG ACACCGAAGC TACTCTCCTT CCAGTCCACA AACGACCAAG 60  
 CCTTCTCTGC CTCAAC ATG GCC AAC AAG GGT CCA TCC TAC GGC ATG AGC 109  
 Met Ala Asn Lys Gly Pro Ser Tyr Gly Met Ser  
 1 5 10  
 25 CGA GAA GTG CAG TCC AAA ATT GAG AAG AAG TAT GAC GAG GAG CTG GAG 157  
 Arg Glu Val Gln Ser Lys Ile Glu Lys Lys Tyr Asp Glu Glu Leu Glu  
 15 20 25  
 30 GAG CGA CTA GTG GAG TGG ATT GTA GTG CAG TGT GGC CCT GAT GTA GGC 205  
 Glu Arg Leu Val Glu Trp Ile Val Val Gln Cys Gly Pro Asp Val Gly  
 30 35 40  
 35 CGC CCA GAT CGT GGG CGC CTG GGC TTC CAG GTG TGG CTG AAG AAT GGT 253  
 Arg Pro Asp Arg Gly Arg Leu Gly Phe Gln Val Trp Leu Lys Asn Gly  
 45 50 55  
 GTG ATT CTG AGC AAA TTG GTG AAC AGC CTG TAT CCT GAG GGA TCG AAG 301  
 Val Ile Leu Ser Lys Leu Val Asn Ser Leu Tyr Pro Glu Gly Ser Lys  
 60 65 70 75  
 40 CCA GTG AAG GTG CCT GAG AAC CCA CCC TCC ATG GTC TTT AAG CAG ATG 349  
 Pro Val Lys Val Pro Glu Asn Pro Pro Ser Met Val Phe Lys Gln Met  
 80 85 90  
 45 GAA CAG GTG GCT CAA TTC TTG AAG GCA GCT GAA GAT TAT GGA GTC ATC 397  
 Glu Gln Val Ala Gln Phe Leu Lys Ala Ala Glu Asp Tyr Gly Val Ile  
 95 100 105  
 50 AAG ACT GAC ATG TTC CAG ACT GTT GAC CTC TAT GAA GGT AAG GAT ATG 445  
 Lys Thr Asp Met Phe Gln Thr Val Asp Leu Tyr Glu Gly Lys Asp Met  
 110 115 120

	GCA GCA GTG CAG AGG ACT CTA ATG GCT TTG GGC AGT TTG GCT GTG ACC	493
	Ala Ala Val Gln Arg Thr Leu Met Ala Leu Gly Ser Leu Ala Val Thr	
	125 130 135	
5	AAA AAC GAT GGA AAC TAC CGT GGA GAT CCC AAC TGG TTT ATG AAG AAA	541
	Lys Asn Asp Gly Asn Tyr Arg Gly Asp Pro Asn Trp Phe Met Lys Lys	
	140 145 150 155	
10	GCC CAG GAG CAT AAG AGG GAC TTC ACA GAC AGC CAA CTG CAG GAG GGG	589
	Ala Gln Glu His Lys Arg Asp Phe Thr Asp Ser Gln Leu Gln Glu Gly	
	160 165 170	
15	AAG CAC GTC ATT GGC CTT CAA ATG GGC AGC AAC AGA GGA GCC TCG CAG	637
	Lys His Val Ile Gly Leu Gln Met Gly Ser Asn Arg Gly Ala Ser Gln	
	175 180 185	
20	GCT GGC ATG ACA GGC TAT GGG CGA CCC CGG CAG ATC ATC AGT TA	681
	Ala Gly Met Thr Gly Tyr Gly Arg Pro Arg Gln Ile Ile Ser	
	190 195 200	
25	GAAAGGGAAG GCCAGCCCTG AGCTGCAGCA TCCTGCTTAG CCTGCCTCAC AAATGCCTAT	741
	GTAGGTTCTT AGCCCTGACA GCTCTGAGGT GTCACTGGGC AAAGATGACT GCACATGGGC	801
30	AGCTCCCACC TATCCTTAGC CTCAGCCCAG CATCTTACCC CAGAGCCACC ACTGCCCTGG	861
	CCCCTGTTCC CAGCTGTACC CCCACCTCTA CTGTTCTCT CATCCTGGAG TAAGCAGGGA	921
35	GAAGTGGGCT GGGGTAGCTG GCTGTAGGCC AGCCCACTGT CCTTGATATC GAATGTCCTT	981
	TGAAGGAGAC CCAGCCCAGC CTCTACATCT TTTCTGGAA TATGTTTTTG GGTGAAATT	1041
	CAAAAAGGAA AAAAGAAAAA TATATAAATA TATATATATA CAAAAA AAAA	1101
		1102

(2) INFORMATION FOR SEQ ID NO: 9:

40 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 201 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

45 (ii) MOLECULE TYPE: protein  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

Met Ala Asn Lys Gly Pro Ser Tyr Gly Met Ser Arg Glu Val Gln Ser
1 5 10 15
Lys Ile Glu Lys Lys Tyr Asp Glu Glu Leu Glu Glu Arg Leu Val Glu
20 25 30



Trp Ile Val Val Gln Cys Gly Pro Asp Val Gly Arg Pro Asp Arg Gly  
35 40 45

5 Arg Leu Gly Phe Gln Val Trp Leu Lys Asn Gly Val Ile Leu Ser Lys  
50 55 60

Leu Val Asn Ser Leu Tyr Pro Glu Gly Ser Lys Pro Val Lys Val Pro  
65 70 75 80

10 Glu Asn Pro Pro Ser Met Val Phe Lys Gln Met Glu Gln Val Ala Gln  
85 90 95

Phe Leu Lys Ala Ala Glu Asp Tyr Gly Val Ile Lys Thr Asp Met Phe  
100 105 110

15 Gln Thr Val Asp Leu Tyr Glu Gly Lys Asp Met Ala Ala Val Gln Arg  
115 120 125

Thr Leu Met Ala Leu Gly Ser Leu Ala Val Thr Lys Asn Asp Gly Asn  
130 135 140

Tyr Arg Gly Asp Pro Asn Trp Phe Met Lys Lys Ala Gln Glu His Lys  
145 150 155 160

25 Arg Asp Phe Thr Asp Ser Gln Leu Gln Glu Gly Lys His Val Ile Gly  
165 170 175

Leu Gln Met Gly Ser Asn Arg Gly Ala Ser Gln Ala Gly Met Thr Gly  
180 185 190

30 Tyr Gly Arg Pro Arg Gln Ile Ile Ser  
195 200

35 (2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 43 base pairs  
(B) TYPE: nucleic acid  
40 (C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

45 ATCGAATTCC GCTACTCTCC TTCCAGCCCA CAAACGACCA AGC

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

ATCAAGCTTG GTGGGAGCTG CCCATGTGCA GTC

33

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

TGCCGTAGGA TGGACCCTTG TTGGC

25

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: modified\_base
- (B) LOCATION:1
- (D) OTHER INFORMATION:/mod\_base= OTHER  
/note= "Y = C or T/U"

(ix) FEATURE:

- (A) NAME/KEY: modified\_base
- (B) LOCATION:4
- (D) OTHER INFORMATION:/mod\_base= OTHER  
/note= "W = A or T/U"

(ix) FEATURE:

- (A) NAME/KEY: modified\_base
- (B) LOCATION:10
- (D) OTHER INFORMATION:/mod\_base= OTHER  
/note= "R = A or G"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

YTAWAAATAR

10

5

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

15

TTTAAATCG

10

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

20

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

TTCAAAATAG

10

30

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 8 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

Cys Cys Cys Met Asn Ser Ser Ser  
1 5

40

45

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 9 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

Lys Arg Gly Gly Cys Lys Arg Arg Lys  
1 5

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 9 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

Thr Lys Asn Asn Gly Asn Ala Ala Lys  
1 5

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

Met Ile Arg Ile Cys Arg Lys Lys  
1 5

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 381 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

AGTCAAGACT AGTTCCACCC AACTCGATTT TAAAGCCTTG CAAGAAGGTG GCTTGTTTGT	60
CCCTTGCAGG TTCCTTTGTC GGGCCAAACT CTAGAATGCC TCCCCCTTC TTTCTCATTG	120
AAGAGCAGAC CCAAGTCCGG GTAACAAGGA AGGGTTTCAG GGTCTGCCC ATAAAAGGTT	180
TTTCCCGGCC GCCCTCAGCA CCGCCCCGCC CCGACCCCCG CAGCATCTCC AAAGCATGCA	240
GAGAATGTCT CCGGCTGCCC CCGACAGACT GCTCCAACCTT GGTGTCTTTC CCCAAATATG	300

GAGCCTGTGT GGAGTGAGTG GGGCGGCCCCG GGGTGGTGAG CCAAGCAGAC TTCCATGGGC 360  
AGGGAGGGGC GCCAGCGGAC G 381

5

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:

10

- (A) LENGTH: 47 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

15

AAGGAAGGGT TTCAGGGTCC TGCCCATAAA AGGTTTTTCC CGGCCGC 47

(2) INFORMATION FOR SEQ ID NO: 22:

20

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 47 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

30

AAGGAAGGGT TTCAGGGTCC TGCCCATAGA TCTTTTTTCC CGGCCGC 47

(2) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:

35

- (A) LENGTH: 43 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

CCGCCCTCAG CACCGCCCCG CCCCAGAGCC CGCAGCATGT CCG 43

45

(2) INFORMATION FOR SEQ ID NO: 24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 43 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

CCGCCCTCAG CACCGCGGAT CCCCACCCC CGCAGCATCT CCG

43

5

(2) INFORMATION FOR SEQ ID NO: 25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 37 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

15

CTCCAAAGCA TGCAGAGAAT GTCTCCGGCT GCCCCCCG

37

(2) INFORMATION FOR SEQ ID NO: 26:

20

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 37 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

CTCGGATCCA TGCTAGCAAT GAATTCGGCT GCCCCCCG

37

30

(2) INFORMATION FOR SEQ ID NO: 27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 44 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

40

TCCAACCTGG TGTCTTTCCC CAAATATGGA GCCTGTGTGG AGTG

44

45

(2) INFORMATION FOR SEQ ID NO: 28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 44 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

TCCAACCTTGG TGTCTTTCCC CAAGGATCCA GCCTGTGTGG AGTG

44

5

(2) INFORMATION FOR SEQ ID NO: 29:

(i) SEQUENCE CHARACTERISTICS:

10

- (A) LENGTH: 44 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

15

TCCAACCTTGG TGTCTTTCCC CGGATATGGA GCCTGTGTGG AGTG

44

(2) INFORMATION FOR SEQ ID NO: 30:

20

(i) SEQUENCE CHARACTERISTICS:

25

- (A) LENGTH: 44 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

30

TCCAACCTTGG TGTCTTTCCC CAAATTAGGA GCCTGTGTGG AGTG

44

(2) INFORMATION FOR SEQ ID NO: 31:

(i) SEQUENCE CHARACTERISTICS:

35

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

40

GGGCAGGGAG GGGCGCCAGC G

21

45 (2) INFORMATION FOR SEQ ID NO: 32:

(i) SEQUENCE CHARACTERISTICS:

50

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

GGGCAGGTAC CGAATTCAGC G

21

5

(2) INFORMATION FOR SEQ ID NO: 33:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 37 base pairs

10

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 33:

15

GGACGGCAGA GGGGTGACAT CACTGCCTAG GCGGCCG

37

(2) INFORMATION FOR SEQ ID NO: 34:

20

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 37 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

25

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34:

GGACGGCAGA GGGGATCCAT GCCTGCCTAG GCGGCCG

37

30

(2) INFORMATION FOR SEQ ID NO: 35:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 37 base pairs

35

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 35:

GGACGGCAGA GGGGATCCAT CACTGCCTAG GCGGCCG

37

40

(2) INFORMATION FOR SEQ ID NO: 36:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 29 base pairs

(B) TYPE: nucleic acid

50

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear



(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 36:

CTGGCTAAAG GGGCGGGGCT TGGCCAGCC

29

5

(2) INFORMATION FOR SEQ ID NO: 37:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 26 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 37:

15

CTCCCATTTC CATGACGTCA TGGTTA

26

(2) INFORMATION FOR SEQ ID NO: 38:

20

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 47 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 38:

AAGGAAGGGT TTCAGGGTCC TGCCCATAGA TCTTTTTTCC CGGCCGC

47

30

(2) INFORMATION FOR SEQ ID NO: 39:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 43 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 39:

40

CCGCCCTCAG CACCGCGGAT CCCCACCCC CGCAGCATCT CCG

43

45

(2) INFORMATION FOR SEQ ID NO: 40:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 37 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 40:

CTCGGATCCA TGCTAGCAAT GAATTCGGCT GCCCCCG

37

5

(2) INFORMATION FOR SEQ ID NO: 41:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 44 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 41:

15

TCCAACCTGG TGTCTTTCCC CAAGGATCCA GCCTGTGTGG AGTG

44

(2) INFORMATION FOR SEQ ID NO: 42:

20

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 44 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 42:

TCCAACCTGG TGTCTTTCCC CGGATATGGA GCCTGTGTGG AGTG

44

30

(2) INFORMATION FOR SEQ ID NO: 43:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 44 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 43:

40

TCCAACCTGG TGTCTTTCCC CAAATTAGGA GCCTGTGTGG AGTG

44

45

(2) INFORMATION FOR SEQ ID NO: 44:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 44:

GGGCAGGTAC CGAATTCAGC G

21

5

(2) INFORMATION FOR SEQ ID NO: 45:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 37 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 45:

15

GGACGGCAGA GGGGATCCAT GCCTGCCTAG GCGGCCG

37

(2) INFORMATION FOR SEQ ID NO: 46:

20

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 37 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 46:

GGACGGCAGA GGGGATCCAT CACTGCCTAG GCGGCCG

37

30

(2) INFORMATION FOR SEQ ID NO: 47:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

35

(ix) FEATURE:

(A) NAME/KEY: modified\_base

(B) LOCATION:3..8

(D) OTHER INFORMATION:/mod\_base= OTHER

/note= "W = A or T"

40

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 47:

CCWWWWWGCC

10

50

(2) INFORMATION FOR SEQ ID NO: 48:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 45 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 48:

CTCCAACCTTG GTGTCTTTCC CCGGATATGG AGCCTGTGTG GAGTG 45

(2) INFORMATION FOR SEQ ID NO: 49:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 45 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 49:

CTCCAACCTTG GTGTCTTTCC CCAAATTAGG AGCCTGTGTG GAGTG 45

(2) INFORMATION FOR SEQ ID NO: 50:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 10 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 50:

CCAAATATGG 10

(2) INFORMATION FOR SEQ ID NO: 51:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 10 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 51:

CCATATATGG 10